Some generalizations of bivariate discrete distributions and their applications

Timothy Opheim and Ram Tripathi
University of Texas at San Antonio
San Antonio, Texas, USA

Presented at the Workshop on DISCRETE DISTRIBUTIONS
IN MEMORY OF PROFESSOR ADRIENNE FREDA KEMP
Saturday, April 13, 2024
Harokopio University of Athens

Abstract

In this paper, we develop a general class of bivariate discrete distributions following the approach of Kundu (2020). Kundu takes each marginal as the random geometric sum of a baseline distribution. We replace the geometric distribution with the more versatile Hurwitz-Lerch-Zeta (HLZ) distribution, which includes, among others, the logseries and the Reimann Zeta as special cases (see Gupta et al. (2008) for more details). Using the logseries as a special case, we develop three specific stopped-sum models: the bivariate Poisson logseries, the bivariate negative binomial logseries and the bivariate binomial logseries. For these models, the joint probability function (pf), the joint moment generating function (mgf), cross moments, method of moment estimators, maximum likelihood estimators, and Fisher information matrices are obtained. Additionally, we identify whether the families are over-dispersed or under-dispersed and develop a result for the distribution of the sum of the components of these bivariate distributions. In order to compare the performance of the MME with the MLE, we obtain asymptotic relative efficiencies for the models considered. Finally, we present a numerical example from the literature.

Key words: Bivariate discrete distribution, generalized Hurwitz-Lerch zeta, method of moments, method of maximum likelihood, information matrix, asymptotic relative efficiency

Applications of Bivariate Discrete Distributions

- In sports, modeling scores of opposing teams (McHale and Scarf 2007, 2011).
- In insurance, modeling number of accidents before and after introduction of regulation, and the number of claims in two different classes (Wu and Yuen 2003).
- Distribution of number of plants of two species in a fixed number of contiguous quadrats in a forest (Gillings 1974).
- Number of accidents incurred by a bus driver in two time periods (Cresswell and Frograt 1963).

Construction of Bivariate Discrete Distributions

- Most constructions of bivariate discrete distributions borrow heavily from the seminal book of Kocherlakota and Kocharlakota (1992) and their review paper in the Encyclopedia of Statistical Sciences published in 2006.
- Lai (2006) broadly categorized the construction of bivariate discrete distributions in 16 clusters.
- Kumar (2008) developed a unified approach to generate bivariate discrete distributions based on hypergeometric factorial moment distributions, called bivariate generalized factorial moment distribution.
- Lee and Cha (2015) utilized minimum and maximum operators to develop two new classes of bivariate discrete distributions.
- Jiang, et al. (2017) developed 18 new distributions by combining various arithmetic operations along with the minimum and maximum operators.
- Kundu (2020) developed a general class of discrete bivariate distributions utilizing the technique of generalizing
- Sellers, Morris and Balakrishnan (2016) and Ong, Gupta, Ma and Sim(2021) developed bivariate version of Comp-Poisson distribution.

Generalizing

- We will use the approach of Kundu (2020) to generate some generalized families of bivariate discrete distributions.
- The technique of generalizing begins by specifying a marginal (baseline) distribution f_N with support {1, 2, … }.
- Let $U_i|N$ and $V_i|N$ be conditionally independent random variables with respective (common) pf's $f_{U|N}$ and $f_{V|N}$ for all $i \in \{1, \ldots, N\}$.
- Let $X = \sum_{i=1}^{N} U_i$, $Y = \sum_{i=1}^{N} V_i$, and denote their conditional pf's given N as $f_{X|N}$ and $f_{Y|N}$, respectively.
- The unconditional distribution of (X, Y) is said to be the resulting generalized bivariate distribution with the joint pf $f_{X,Y}$.

Outline of Presentation

- Introduce the Hurwitz-Lerch Zeta (HLZ), general family which includes many well-known distributions as special cases: Log series, Lotka, Reimann Zeta, Good, Estoup, Zipf Mandelbrot (See Gupta et al. (2008)
- An example where the baseline distribution is a Hurwitz-Lerch Zeta (HLZ) distribution and the conditional distributions are Poisson, Negative Binomial and Binomial (Following the approach in a recent paper by Liew et al. (2020)).
- An example where the baseline distribution is a logarithmic distribution and the conditional distributions are Poisson, Negative Binomial, Binomial.
- In all instances, means, variances, and the covariance are derived.
- Statistical inference for cases with a baseline logarithmic model are provided along with a comparison of their MLEs and MMEs via their multivariate asymptotic relative efficiency (ARE).
- Provide an example based on the Italian Series A football data; obtain parameter estimates for the bivariate Poisson and bivariate negative binomial models.

Polylogarithm and Pochhammer symbol

- The preceding distributions do not have a closed-form representation since the Lerch transcendent function admits of a closed-form representation only when $s \in \{-1, 0, \dots\}$.
- A special case of the Lerch transcendent function is the polylogarithm $\mathrm{Li}_s(z) = \sum_{n=1}^{\infty} n^{-s} z^n$, where $s \in \mathbb{R}$ and $z \in (-1,1)$.
- Let $S(n, k)$ denote a Stirling number of the second kind.
- Two pertinent closed-form representations of the polylogarithm are

$$\mathrm{Li}_1(z) = -\ln(1-z)$$

$$\mathrm{Li}_{-s}(z) = \sum_{k=0}^{s} k!\, S(s+1, k+1) \left(\frac{z}{1-z}\right)^{k+1}, \ s \in \{0, 1, \dots\}.$$

Polylogarithm and Pochhammer symbol

- The Pochhammer symbol is defined as $(x)_n = \left(\Gamma(x)\right)^{-1}\Gamma(x+n)$.
- Let $s(n,k)$ denote a Stirling number of the first kind. When n is a non-negative integer, a closed-form representation of $(x)_n$ is $\sum_{k=0}^{n}(-1)^{n-k}s(n,k)x^k$.
- Let ψ_0 and ψ_1 represent the digamma and trigamma functions, respectively. The following partials are useful for likelihood calculations.

$$\frac{\partial \mathrm{Li}_s(z)}{\partial z} = z^{-1}\mathrm{Li}_{s-1}(z)$$

$$\frac{\partial (x)_n}{\partial x} = (x)_n[\psi_0(x+n) - \psi_0(x)]$$

$$\frac{\partial \psi_0(x)}{\partial x} = \psi_1(x)$$

Baseline HLZ Distribution

- Let $N \sim HLZ(p, s, a)$ with pf

$$f_N(n) = \frac{1}{p\Phi(p,s+1,a+1)} \frac{p^n}{(n+a)^{s+1}},$$

where $\Phi(z, s, a) = \sum_{k=0}^{\infty} \frac{z^k}{(k+a)^s}$ is the Lerch transcendent and the parameter space is defined by the intersection of the space: $p \in (0,1), s \in \mathbb{R}, a > -1$.

- Using the previous definitions on X and Y,

$$f_{X,Y}(x, y) = \frac{1}{p\Phi(p, s+1, a+1)} \sum_{n=1}^{\infty} \frac{p^n}{(n+a)^{s+1}} f_{X|N}(x|n) f_{Y|N}(y|n)$$

Marginal pdf's are:

$$f_X(x) = \frac{1}{p\Phi(p,s+1,a+1)} \sum_{n=1}^{\infty} \frac{p^n}{(n+a)^{s+1}} f_{X|N}(x|n)$$

$$f_Y(y) = \frac{1}{p\Phi(p,s+1,a+1)} \sum_{n=1}^{\infty} \frac{p^n}{(n+a)^{s+1}} f_{Y|N}(y|n)$$

Baseline HLZ Distribution

- Moments

$$E[X] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)E(U|N)$$

$$E[Y] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)E(V|N)$$

$$Var[X] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)Var(U|N) + \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)E^2[U|N]$$

$$Var[Y] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)Var(V|N) + \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)E^2[V|N]$$

$$Cov[X,Y] = \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\mathrm{E[U|N]E[V|N]}$$

Baseline HLZ Distribution (Poisson)

- Let $U_i|N \sim Poi(\lambda_1)$ and $V_i|N\sim Poi(\lambda_2)$ be conditionally independent random variables. Hence, $X|N \sim Poi(N\lambda_1)$ and $Y|N \sim Poi(N\lambda_2)$.
- The unconditional joint distribution of (X, Y) is

$$f_{X,Y}(x,y) = \frac{\lambda_1^x \lambda_2^y e^{-(\lambda_1+\lambda_2)}}{x!y!\Phi(p,s+1,a+1)} \sum_{i=0}^{x+y} \binom{x+y}{i} (-a)^{x+y-i}\, \Phi\left(pe^{-(\lambda_1+\lambda_2)}, s+1-i, a+1\right).$$

- The unconditional marginal distributions are

$$f_X(x) = \frac{\lambda_1^x e^{-\lambda_1}}{x!\Phi(p,s+1,a+1)} \sum_{i=0}^{x} \binom{x}{i} (-a)^{x-i}\, \Phi\left(pe^{-\lambda_1}, s+1-i, a+1\right)$$

$$f_Y(y) = \frac{\lambda_2^y e^{-\lambda_2}}{y!\Phi(p,s+1,a+1)} \sum_{i=0}^{y} \binom{y}{i} (-a)^{y-i}\, \Phi\left(pe^{-\lambda_2}, s+1-i, a+1\right)$$

Baseline HLZ Distribution (Poisson)

- **Moments**

$$E[X] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\lambda_1$$

$$E[Y] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\lambda_2$$

$$Var[X] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\lambda_1 + \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\lambda_1^2$$

$$Var[Y] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\lambda_2 + \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\lambda_2^2$$

$$Cov[X,Y] = \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\lambda_1\lambda_2$$

Baseline HLZ Distribution (Negative Binomial)

- Let $U_i|N \sim NB(r_1, \theta_1)$ and $V_i|N \sim NB(r_2, \theta_2)$ be conditionally independent random variables. Hence, $X|N \sim NB(Nr_1, \theta_1)$and $Y|N \sim NB(Nr_2, \theta_2)$.
- The unconditional joint distribution of (X, Y) is

$$f_{X,Y}(x,y) = \frac{\theta_1^x \theta_2^y (-1)^{x+y}}{x!y!\Phi(p,s+1,a+1)} \sum_{i=0}^{x} \sum_{j=0}^{y} s(x,i)s(y,j)r_1^i r_2^j \sum_{k=0}^{i+j} \binom{i+j}{k} (-1)^k a^{i+j-k} \Phi(p(1-\theta_1)^{r_1}(1-\theta_2)^{r_2}, s+1-k, a+1).$$

- The unconditional marginal distributions are

$$f_X(x) = \frac{\theta_1^x (-1)^x}{x!\ \Phi(p, s+1, a+1)} \sum_{i=0}^{x} s(x,i) r_1^i \sum_{k=0}^{i} \binom{i}{k} (-1)^k a^{i-k} \Phi(p(1-\theta_1)^{r_1}, s+1-k, a+1)$$

$$f_Y(y) = \frac{\theta_y^2 (-1)^y}{\mathrm{y}!\ \Phi(p, s+1, a+1)} \sum_{i=0}^{y} s(y,i) r_2^i \sum_{k=0}^{i} \binom{i}{k} (-1)^k a^{i-k} \Phi(p(1-\theta_2)^{r_2}, s+1-k, a+1)\ .$$

Baseline HLZ Distribution (Negative Binomial)

- **Moments**

$$E[X] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\frac{r_1\theta_1}{1-\theta_1}$$

$$E[Y] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\frac{r_2\theta_2}{1-\theta_2}$$

$$Var[X] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\frac{r_1\theta_1}{(1-\theta_1)^2} + \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\left(\frac{r_1\theta_1}{1-\theta_1}\right)^2$$

$$Var[Y] = \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)\frac{r_2\theta_2}{(1-\theta_2)^2} + \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\left(\frac{r_2\theta_2}{1-\theta_2}\right)^2$$

$$Cov[X,Y] = \left(\frac{\Phi(p,s-1,a+1)}{\Phi(p,s+1,a+1)} - \left(\frac{\Phi(p,s,a+1)}{\Phi(p,s+1,a+1)} - a\right)^2\right)\frac{r_1\theta_1}{1-\theta_1}\frac{r_2\theta_2}{1-\theta_2}$$

Special Case: Baseline Logarithmic Distribution

- Let $a = -(\ln(1-p))^{-1}$
- Let $N \sim Log(p)$ with pf $f_N(n) = an^{-1}p^n$, where $p \in (0,1)$.
- The unconditional joint distribution of X and Y is

$$f_{X,Y}(x,y) = a \sum_{n=1}^{\infty} \frac{p^n}{n} f_{X|N}(x|n) f_{Y|N}(y|n).$$

- The marginal distributions of X and Y are

$$f_X(x) = a \sum_{n=1}^{\infty} \frac{p^n}{n} f_{X|N}(x|n)$$

$$f_Y(y) = a \sum_{n=1}^{\infty} \frac{p^n}{n} f_{Y|N}(y|n)$$

Baseline Logarithmic Distribution

- **Moments**

$E[X] = ap(1-p)^{-1}E(U|N)$
$E[Y] = ap(1-p)^{-1}E(V|N)$
$Var[X] = ap(1-p)^{-1}(Var(U|N) + (1-ap)(1-p)^{-1}E^2[U|N])$
$Var[Y] = ap(1-p)^{-1}(Var(V|N) + (1-ap)(1-p)^{-1}E^2[V|N])$
$Cov[X,Y] = ap(1-ap)(1-p)^{-2}\mathrm{E[U|N]E[V|N]}$

- From the above moments, regardless of the distributions $U|N$ and $V|N$, one possible MME for p is $g^{-1}\left(\left(1+\frac{n-1}{n}\frac{s_{xy}}{\bar{x}\bar{y}}\right)^{-1}\right)$, where $g(p) = -p\,(\ln(1-p))^{-1}$.

Comments on MME of $\boldsymbol{p}$

- $\tilde{p} = g^{-1}\left(\left(1 + \frac{n-1}{n}\frac{s_{xy}}{\bar{x}\bar{y}}\right)^{-1}\right)$
- Although $g(p) = -p\,(\ln(1-p))^{-1}$ is a strictly decreasing function over the domain $(0,1)$ with a range of $(0,1)$, no closed-form expression exists for its inverse; hence, an iterative procedure is necessary.
- Moreover, when s_{xy} is negative, this MME of p is outside of the parameter space of p, thus requiring employment of another estimator.

Baseline Logarithmic Distribution (Poisson)

- Let $U_i|N \sim Poi(\lambda_1)$ and $V_i|N\sim Poi(\lambda_2)$ be conditionally independent random variables for all i. Hence, $X|N \sim Poi(N\lambda_1)$ and $Y|N \sim Poi(N\lambda_2)$.
- Let $\boldsymbol{\theta}_P = (\lambda_1, \lambda_2, p)$ and $z_P = z_P(\boldsymbol{\theta}_P) = pe^{-(\lambda_1+\lambda_2)}$.
- The unconditional joint distribution of (X, Y) is

$$f_{X,Y}(x,y) = a\frac{\lambda_1^x}{x!}\frac{\lambda_2^y}{y!}\mathrm{Li}_{1-(x+y)}(z_P) = \begin{cases} -a\ln(1-z_P), & if\ (x,y)=(0,0) \\ \mathrm{a}\frac{\lambda_1^x}{x!}\frac{\lambda_2^y}{y!}\sum_{i=0}^{x+y-1} i!\,S(x+y,i+1)\left(\frac{z_P}{1-z_P}\right)^i, & o.w. \end{cases}$$

Baseline Logarithmic Distribution (Poisson)

- **Moments**

$E[X] = ap(1-p)^{-1}\lambda_1$
$E[Y] = ap(1-p)^{-1}\lambda_2$
$Var[X] = ap(1-p)^{-1}\lambda_1(1+(1-ap)(1-p)^{-1}\lambda_1)$
$Var[Y] = ap(1-p)^{-1}\lambda_2(1+(1-ap)(1-p)^{-1}\lambda_2)$
$Cov[X,Y] = ap(1-ap)(1-p)^{-2}\lambda_1\lambda_2$

- Since $(1-ap)(1-p)^{-1} > 0$ for all $p \in (0,1)$, X and Y are overdispersed.
- Using the MME of p, the MMEs of the remaining parameters are

$$\tilde{\lambda}_1 = (1-\tilde{p})(\tilde{a}\tilde{p})^{-1}\bar{x}$$
$$\tilde{\lambda}_2 = (1-\tilde{p})(\tilde{a}\tilde{p})^{-1}\bar{y}$$

Baseline Logarithmic Distribution (Poisson)

- We now develop the partial derivatives of the log likelihood (based on a sample size of m pairs) as well as the Fisher information matrix. The following notation is employed

$$A_0 = \sum_{i=1}^{m} \ln\left[\mathrm{Li}_{1-(x_i+y_i)}(z_P)\right] \qquad A_1 = \sum_{i=1}^{m} \frac{\mathrm{Li}_{-(x_i+y_i)}(z_P)}{\mathrm{Li}_{1-(x_i+y_i)}(z_P)}$$

$$A_2 = \sum_{i=1}^{m} \frac{\mathrm{Li}_{-(x_i+y_i+1)}(z_P)}{\mathrm{Li}_{1-(x_i+y_i)}(z_P)} \qquad A_3 = \sum_{i=1}^{m} \left(\frac{\mathrm{Li}_{-(x_i+y_i)}(z_P)}{\mathrm{Li}_{1-(x_i+y_i)}(z_P)}\right)^2$$

$$a_1 = \mathrm{E}_{X,Y}\left[\frac{\mathrm{Li}_{-(X+Y)}(z_P)}{\mathrm{Li}_{1-(X+Y)}(z_P)}\right] \qquad a_2 = \mathrm{E}_{X,Y}\left[\frac{\mathrm{Li}_{-(X+Y)}(z_P)}{\mathrm{Li}_{1-(X+Y)}(z_P)} - \left(\frac{\mathrm{Li}_{-(X+Y)}(z_P)}{\mathrm{Li}_{1-(X+Y)}(z_P)}\right)^2\right]$$

- For a given $\boldsymbol{\theta}_P$, the expectations can be approximated by truncating the infinite series or by Monte Carlo methods.

Baseline Logarithmic Distribution (Poisson)

- Up to an additive constant, the log-likelihood of a sample size m from the bivariate Poisson logarithmic distribution is

$$l(\boldsymbol{\theta}_P) = m\bar{x}\ln(\lambda_1) + m\bar{y}\ln(\lambda_2) - m\ln(-\ln(1-p)) + A_0$$

- Note

$$\mathrm{d}A_0 = \mathrm{A}_1(p^{-1}\mathrm{d}p - \mathrm{d}\lambda_1 - \mathrm{d}\lambda_2) \text{ and}$$

$$\mathrm{d}A_1 = (A_2 - A_3)(p^{-1}\mathrm{d}p - \mathrm{d}\lambda_1 - \mathrm{d}\lambda_2).$$

Hence the first two differentials of the log-likelihood function are

$$\mathrm{d}l(\boldsymbol{\theta}_P) = (m\bar{x}\lambda_1^{-1} - A_1)\mathrm{d}\lambda_1 + (m\bar{y}\lambda_2^{-1} - A_1)\mathrm{d}\lambda_2 + (m[(1-p)\ln(1-p)]^{-1} + p^{-1}A_1)\mathrm{d}p$$

$$\mathrm{d}^2 l(\boldsymbol{\theta}_P) = (-m\bar{x}\lambda_1^{-2} + A_2 - A_3)(\mathrm{d}\lambda_1)^2 + (-m\bar{y}\lambda_2^{-2} + A_2 - A_3)(\mathrm{d}\lambda_2)^2 + \big(ma(a-1)(1-p)^{-2} + p^{-2}(A_2 - A_1 - A_3)\big)(\mathrm{d}p)^2 + 2(A_2 - A_3)\mathrm{d}\lambda_1\mathrm{d}\lambda_2 - 2(A_2 - A_3)p^{-1}\mathrm{d}\lambda_1\mathrm{d}p - 2(A_2 - A_3)p^{-1}\mathrm{d}\lambda_2\mathrm{d}p$$

Baseline Logarithmic Distribution (Poisson)

- Let $f_1 = ap(1-p)^{-1}, f_2 = a(1-a)(1-p)^{-2}$, and $a^* = a_1 - a_2$.
- Taking the negative expectation of the Hessian matrix, the information matrix is

$$I_m(\boldsymbol{\theta}_P) = \begin{pmatrix} f_1\lambda_1^{-1} - a_2 & -a_2 & -a_2 p^{-1} \\ \cdot & f_1\lambda_2^{-1} - a_2 & -a_2 p^{-1} \\ \cdot & \cdot & f_2 + a^* p^{-2} \end{pmatrix},$$

where the missing values are filled in by symmetry.

Baseline Logarithmic Distribution (Binomial)

- Let $U_i|N \sim B(n_1, p_1)$ and $V_i|N \sim B(n_2, p_2)$ be conditionally independent random variables for all i with n_1 and n_2 known. Hence, $X|N \sim B(n_1 N, p_1)$and $Y|N \sim B(n_2 N, p_2)$.
- Let $\boldsymbol{\theta}_B = (p_1, p_2, p)$ and $z_B = z_B(\boldsymbol{\theta}_B) = p(1-p_1)^{n_1}(1-p_2)^{n_2}$.
- A useful expression of the unconditional joint distribution of (X, Y) for statistical inference is

$$f_{X,Y}(x,y) = \frac{a}{x!\,y!}\left(\frac{p_1}{1-p_1}\right)^{n_1}\left(\frac{p_2}{1-p_2}\right)^{n_2}\sum_{n=1}^{\infty}\frac{1}{n}(nn_1 - x + 1)_x (nn_2 - y + 1)_y z_B^n$$

- A closed-form representation of the above density is

$$f_{X,Y}(x,y) = \frac{a}{x!\,y!}\left(\frac{p_1}{1-p_1}\right)^{n_1}\left(\frac{p_2}{1-p_2}\right)^{n_2}\sum_{i=0}^{x}\sum_{j=0}^{y}\sum_{k=0}^{i}\sum_{l=0}^{i}\binom{i}{k}\binom{j}{l}(-1)^{x+y-(i+j)}$$

$$\times\, s(x,i)s(y,j)n_1^k n_2^l (1-x)^{i-k}(1-y)^{j-l}\mathrm{Li}_{1-(k+l)}(z_B)$$

Baseline Logarithmic Distribution (Binomial)

- **Moments**

$E[X] = ap(1-p)^{-1}n_1p_1$
$E[Y] = ap(1-p)^{-1}n_2p_2$
$Var[X] = ap(1-p)^{-1}n_1p_1(1-p_1+(1-ap)(1-p)^{-1}n_1p_1)$
$Var[Y] = ap(1-p)^{-1}n_2p_2(1-p_2+(1-ap)(1-p)^{-1}n_2p_2)$
$Cov[X,Y] = ap(1-ap)(1-p)^{-2}n_1p_1n_2p_2$

- Clearly $n_1 > (<)(1-p)(1-ap)^{-1}$ implies X is overdispersed (underdispersed). A similar conclusion holding for Y.

- Using the MME of p, the MMEs of the remaining parameters are

$$\tilde{p}_1 = (1-\tilde{p})(n_1\tilde{a}\tilde{p})^{-1}\bar{x}$$
$$\tilde{p}_2 = (1-\tilde{p})(n_2\tilde{a}\tilde{p})^{-1}\bar{y}$$

Baseline Logarithmic Distribution (Binomial)

- As before, simplifying notation is useful for the ensuing statistical inference

$$B_0 = \sum_{i=1}^{m} \ln\left[\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k\right]$$

$$B_1 = \sum_{i=1}^{m} \frac{\sum_{k=1}^{\infty}(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k}{\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k}$$

$$B_2 = \sum_{i=1}^{m} \frac{\sum_{k=1}^{\infty} k(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k}{\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k}$$

$$B_3 = \sum_{i=1}^{m} \left(\frac{\sum_{k=1}^{\infty}(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k}{\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - x + 1)_x (kn_2 - y + 1)_y z_B^k}\right)^2$$

$$b_1 = \mathrm{E}_{X,Y}\left[\frac{\sum_{k=1}^{\infty}(kn_1 - X + 1)_X (kn_2 - Y + 1)_Y z_B^k}{\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - X + 1)_X (kn_2 - Y + 1)_Y z_B^k}\right]$$

$$b_2 = \mathrm{E}_{X,Y}\left[\left(\frac{\sum_{k=1}^{\infty}(kn_1 - X + 1)_X (kn_2 - Y + 1)_Y z_B^k}{\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - X + 1)_X (kn_2 - Y + 1)_Y z_B^k}\right)^2 - \frac{\sum_{k=1}^{\infty} k(kn_1 - X + 1)_X (kn_2 - Y + 1)_Y z_B^k}{\sum_{k=1}^{\infty} \frac{1}{k}(kn_1 - X + 1)_X (kn_2 - Y + 1)_Y z_B^k}\right]$$

Baseline Logarithmic Distribution (Binomial)

- Up to an additive constant, the log-likelihood of a sample size m from the bivariate Binomial logarithmic distribution is

$$l(\boldsymbol{\theta}_B) = m\bar{x}\ln\left(\frac{p_1}{1-p_1}\right) + m\bar{y}\ln\left(\frac{p_2}{1-p_2}\right) - m\ln(-\ln(1-p)) + B_0$$

Note, $\mathrm{d}B_0 = B_1(p^{-1}\mathrm{d}p - n_1(1-p_1)^{-1}\mathrm{d}p_1 - n_2(1-p_2)^{-1}\mathrm{d}p_2)$ and

$$\mathrm{d}B_1 = (B_2 - B_3)(p^{-1}\mathrm{d}p - n_1(1-p_1)^{-1}\mathrm{d}p_1 - n_2(1-p_2)^{-1}\mathrm{d}p_2)$$

The first two differentials of the log-likelihood function are

$$\mathrm{d}l(\boldsymbol{\theta}_B) = (m\bar{x}[p_1(1-p_1)]^{-1} - n_1(1-p_1)^{-1}B_1)\mathrm{d}p_1 + (m\bar{y}[p_2(1-p_2)]^{-1} - n_2(1-p_2)^{-1}B_1)\mathrm{d}p_2 + (m[(1-p)\ln(1-p)]^{-1} + p^{-1}B_1)\mathrm{d}p$$

$$\mathrm{d}^2l(\boldsymbol{\theta}_B) = -(m\bar{x}(1-2p_1)[p_1(1-p_1)]^{-2} + n_1(1-p_1)^{-2}(B_1 + n_1(B_3 - B_2)))(\mathrm{d}p_1)^2 - (m\bar{y}(1-2p_2)[p_2(1-p_2)]^{-2} + n_2(1-p_2)^{-2}(B_1 + n_2(B_3 - B_2)))(\mathrm{d}p_2)^2 - \left(ma(1-a)(1-p)^{-2} + p^{-2}(B_1 + B_3 - B_2)\right)(\mathrm{d}p)^2 - 2n_1(1-p_1)^{-1}n_2(1-p_2)^{-1}(B_3 - B_2)\mathrm{d}p_1\,\mathrm{d}p_2 + 2n_1(1-p_1)^{-1}p^{-1}(B_3 - B_2)\mathrm{d}p_1\,\mathrm{d}p + 2n_2(1-p_2)^{-1}p^{-1}(B_3 - B_2)\mathrm{d}p_2\,\mathrm{d}p$$

Baseline Logarithmic Distribution (Binomial)

- Let $b_1^* = b_1 + n_1 b_2, b_2^* = b_1 + n_2 b_2$, and $b_3^* = b_1 + b_2$.

- Taking the negative expectation of the Hessian matrix, the information matrix is

$$I_m(\boldsymbol{\theta}_P) = \begin{pmatrix} \frac{n_1}{(1-p_1)^2}\left[f_1 \frac{1-2p_1}{p_1} + b_1^*\right] & b_2 \frac{n_1 n_2}{(1-p_1)(1-p_2)} & b_2 \frac{n_1}{p(1-p_1)} \\ . & \frac{n_2}{(1-p_2)^2}\left[f_1 \frac{1-2p_2}{p_2} + b_2^*\right] & b_2 \frac{n_2}{p(1-p_2)} \\ . & . & f_2 + b_3^* p^{-2} \end{pmatrix},$$

where the missing values are filled in by symmetry.

Baseline Logarithmic Distribution (Negative Binomial)

- Let $U_i|N \sim NB(r_1, p_1)$ and $V_i|N \sim NB(r_2, p_2)$ be conditionally independent random variables for all i. Hence, $X|N \sim NB(r_1 N, p_1)$ and $Y|N \sim NB(r_2 N, p_2)$.
- Let $\boldsymbol{\theta}_{NB} = (r_1, r_2, p_1, p_2, p)$ and $z_{NB} = z_{NB}(\boldsymbol{\theta}_{NB}) = p(1-p_1)^{r_1}(1-p_2)^{r_2}$.
- A useful expression of the unconditional joint distribution of (X, Y) for statistical inference is

$$f_{X,Y}(x,y) = a\frac{p_1^x p_2^y}{x!\,y!}\sum_{n=1}^{\infty}\frac{1}{n}(nr_1)_x(nr_2)_y z_{NB}^n$$

- A closed-form representation of the above density is

$$f_{X,Y}(x,y) = a\frac{p_1^x p_2^y}{x!\,y!}\sum_{i=0}^{x}\sum_{j=0}^{y}(-1)^{x+y-(i+j)}r_1^i r_2^j s(x,i)s(y,j)\mathrm{Li}_{1-(i+j)}(z_{NB})$$

Baseline Logarithmic Distribution (Negative Binomial)

- **Moments**

$$E[X] = ap(1-p)^{-1}r_1p_1(1-p_1)^{-1}$$
$$E[Y] = ap(1-p)^{-1}r_2p_2(1-p_2)^{-1}$$
$$Var[X] = ap(1-p)^{-1}r_1p_1(1-p_1)^{-2}(1+(1-ap)(1-p)^{-1}r_1p_1)$$
$$Var[Y] = ap(1-p)^{-1}r_2p_2(1-p_2)^{-2}(1+(1-ap)(1-p)^{-1}r_2p_2)$$
$$Cov[X,Y] = ap(1-ap)(1-p)^{-2}r_1p_1r_2p_2(1-p_1)^{-1}(1-p_2)^{-1}$$

- Since $(1-ap)(1-p)^{-1} > 0$ for all $p \in (0,1)$ and $r_1, r_2 > 0$, X and Y are overdispersed.

Baseline Logarithmic Distribution (Negative Binomial)

- Using the MME of p, the MMEs of the remaining parameters are

$$\tilde{r}_1 = -\frac{(1-\tilde{p})(1-\tilde{p}_1)\ln(1-\tilde{p})}{\tilde{p}\tilde{p}_1}\bar{x}$$
$$\tilde{r}_2 = -\frac{(1-\tilde{p})(1-\tilde{p}_2)\ln(1-\tilde{p})}{\tilde{p}\tilde{p}_2}\bar{y}$$
$$\tilde{p}_1 = 1 - \left[\frac{n-1}{n}\frac{s_x^2}{\bar{x}} + \left(\frac{\ln(1-\tilde{p})}{\tilde{p}} + 1\right)\bar{x}\right]^{-1}$$
$$\tilde{p}_2 = 1 - \left[\frac{n-1}{n}\frac{s_y^2}{\bar{y}} + \left(\frac{\ln(1-\tilde{p})}{\tilde{p}} + 1\right)\bar{y}\right]^{-1}$$

- Using the approach in the previous two cases, the likelihood and its partials may be obtained. However, these formulas are not shown for constraints on space.

General Comments

- In each of these bivariate cases, the usual limiting properties hold.
- MLEs require more computing power than MMEs to calculate, but the difference is practically negligible.
- For verification of the practitioner’s coding efforts, view our paper for an example.

ARE IN MULTIPARAMETER CASE

- Let $\tilde{\boldsymbol{\theta}}$ denote a vector (of length p) of MMEs.
- The multivariate asymptotic relative efficiency is defined as

$$\text{ARE}_p = \left(\frac{|I_m(\boldsymbol{\theta})|}{\left|\text{Var}\left(\tilde{\boldsymbol{\theta}}\right)\right|} \right)^{\frac{1}{p}}.$$

- And an estimator of it is

$$\widehat{\text{ARE}}_p = \left(\frac{|I_m(\boldsymbol{\theta})|}{\left|\widehat{\text{Var}}\left(\tilde{\boldsymbol{\theta}}\right)\right|} \right)^{\frac{1}{p}}.$$

Biv. Poisson Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations

λ_1 \ λ_2	n	p = .5					p = .7				
		1	2	3	4	5	1	2	3	4	5
1	50	.88	.81	.76	.71	.67	.67	.64	.59	.55	.52
	100	.83	.79	.74	.70	.66	.70	.65	.61	.57	.53
2	50	.81	.79	.76	.72	.69	.63	.61	.59	.55	.52
	100	.79	.78	.75	.72	.68	.65	.63	.59	.56	.63
3	50	.76	.76	.74	.71	.68	.59	.58	.56	.53	.51
	100	.74	.75	.73	.70	.67	.60	.59	.57	.54	.51
4	50	.71	.73	.71	.69	.66	.55	.55	.53		
	100	.70	.71	.70	.67	.65	.56	.56	.54	.51	
5	50	.67	.69	.68	.66	.64	.52	.52	.51	.49	
	100	.66	.68	.67	.65	.62	.53	.53	.51	.49	

Biv. Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $n_1 = 10$ and $n_2 = 10$

p_1 \ p_2	n	p = .5			p = .7		
		.3	.5	.7	.3	.5	.7
.3	50	.65	.52	.37	.47	.37	.25
	100	.63	.51	.35	.48	.37	.25
.5	50	.52	.45	.35	.37	.31	.23
	100	.51	.44	.34	.37	.31	.23
.7	50	.36	.34	.30	.25	.23	.19
	100	.35	.33	.29	.25	.23	.20

Biv. Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $n_1 = 15$ and $n_2 = 15$

p_1 \ p_2	n	p = .5			p = .7		
		.3	.5	.7	.3	.5	.7
.3	50	.57	.44	.31	.40	.30	.21
	100	.55	.42	.30	.40	.30	.21
.5	50	.44	.37	.29	.30	.25	.20
	100	.43	.36	.29	.30	.25	.20
.7	50	.31	.29	.26	.21	.20	.19
	100	.30	.29	.26	.21	.20	.20

Biv. Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $n_1 = 10$ and $n_2 = 15$

p_1 \ p_2	n	p = .5			p = .7		
		.3	.5	.7	.3	.5	.7
.3	50	.60	.45	.31	.43	.31	.21
	100	.59	.44	.30	.44	.31	.21
.5	50	.50	.40	.30	.34	.27	.20
	100	.49	.39	.29	.35	.27	.20
.7	50	.36	.33	.28	.24	.22	.19
	100	.35	.32	.27	.24	.22	.19

Biv. Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $n_1 = 15$ and $n_2 = 10$

p_1 \ p_2	n	p = .5			p = .7		
		.3	.5	.7	.3	.5	.7
.3	50	.60	.50	.36	.43	.35	.24
	100	.59	.48	.35	.44	.35	.24
.5	50	.45	.40	.30	.31	.27	.22
	100	.44	.39	.31	.31	.27	.22
.7	50	.31	.30	.28	.21	.20	.19
	100	.30	.29	.27	.21	.20	.19

Biv. Negative Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $r_1 = 3, r_2 = 3$, and $n = 1000$.

p_1 \ p_2	p = .5			p = .7		
	.3	.5	.7	.3	.5	.7
.3	.05	.19	.29	.003	.01	.02
.5	.24	.41	.38	.02	.05	.06
.7	.29	.37	.33	.02	.05	.07

Biv. Negative Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $r_1 = 1, r_2 = 1$, and $n = 1000$.

p_1 \ p_2	p = .5			p = .7		
	.3	.5	.7	.3	.5	.7
.3	.12	.41	.52	.02	.14	.28
.5	.51	.66	.67	.25	.45	.48
.7	.52	.66	.66	.28	.47	.48

Biv. Negative Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $r_1 = 1, r_2 = 3$, and $n = 1000$.

p_1 \ p_2	p = .5			p = .7		
	.3	.5	.7	.3	.5	.7
.3	.05	.12	.13	.002	.01	.005
.5	.19	.27	.31	.04	.05	.03
.7	.07	.38	.40	.05	.06	.03

Biv. Negative Binomial Logarithmic (MMEs vs. MLEs)

- Estimated multivariate asymptotic relative efficiency (ARE) of various parameter combinations when $r_1 = 3, r_2 = 1$, and $n = 1000$.

p_1 \ p_2	p = .5			p = .7		
	.3	.5	.7	.3	.5	.7
.3	.03	.15	.17	.02	.03	.04
.5	.06	.18	.33	.02	.01	.05
.7	.08	.13	.31	.002	.02	.02

Example Data: Italian Football Series A data Between "ACF Firontina" (X) and "Juventus" (Y) during 1990 to 2005

Obs.	ACF Firontina(X)	Juventus (Y)	Obs.	ACF Firontina(X	Juventus (Y)
1	1	2	14	4	1
2	0	0	15	4	4
3	0	0	16	1	3
4	2	2	17	1	3
5	4	3	18	0	0
6	0	1	19	1	0
7	1	0	20	0	2
8	3	2	21	3	0
9	1	3	22	3	0
10	2	0	23	1	2
11	1	2	24	1	4
12	2	3	25	0	2
13	0	0	26	0	5

Example Contd.

For this data:

n=26, $\bar{x}$ =1.385, s_x^2=1.846, $\overline{y}$ =1.692, s_y^2=2.302, s_{xy}=0.243

The MME's of the parameters for two of the models for this data are:

1) **Bivariate Poisson Log series Model:**

$\tilde{p}$ = .176[±.0522], $\tilde{\lambda}_1$= 1.255[±.0694], $\tilde{\lambda}_2$= 1.534[±.0857]

2) For the Negative Binomial Log series Model:

$\tilde{\theta}_{1,NegBin}$ = .126[±.0353]

$\tilde{\theta}_{2,NegBin}$= .122[±0426]

$\tilde{r}_1$= 6.297[±4.5284]

$\tilde{r}_2$= 6.526 [±5.5867]

References

1. Cresswell, W. L. and Froggatt, P. (1963). The causation of bus driver acceidents,

An epidemiological study. it Oxford University Press, London.

2. Gupta, Pushpa L, Gupta, Ramesh C., Seng-Huat, Ong, and Srivastava, H. M. (2008). A class of Hurwitz-Lerch Zeta distributions and their applications in reliability. Applied Mathematics and Computation, 196, pp. 521-531.

3. Gillings,, D. B. (1974). Some further results for bivariate generalizations of the Neymann

Type A distributuion. Biometrics, 30, pp. 619-628.

3. Jiang, X., Chu, J. and Nadarajah, S. (2017). New classes of discrete bivariate distributions with applications to football data. Communications in Statistics, Theory and Methods, 46, pp. 8069-8085.

4. Kocherlakota, S. and Kocherlakota, K. (1992). Bivariate discrete distributions, Marcell and Dekker Inc., New York.

5. Kocherlakota, S. and Kocherlakota, K. (2006). Bivariate discrete distributions. Encyclopedia of Statistical Sciences, Samuel Kotz, Campbell Read, N. Balakrishnan, Brani Vidakovic and and Norman L. Johnson Editors. John Wiley and Sons.

6. Kumar, C. Satheesh (2008). A unied approach to bivariate discrete distributions. Metrika, 67, pp. 113-123.

7. Kundu, Debasis (2020). On a general class of discrete bivariate distributions. Sankhya, B,82, pp. 270-304.

References Contd.

8. Lai, C. D. (2006). Construction of discrete bivariate distributions. Advances in Disribution Theory, Order Statistics and Inference, (Eds. N. Balakrishnan, E. Castillo, and J. M. Sarabia), Birkhauser, Boston.

9. Lee, Hyunju and Cha, Ji Huan (2015). On two general classes of discrete bivariate, distributions. The American Statistician, 69, pp. 221-230. Liew, Kian W., Ong, Seng H. and Toh, Kian K. (2020). The Poisson-stopped Hurwitz-Lerch zeta distribution. To appear in Communications in Statistics, Theory and Methods.

10. McHale, I. and Scarf, P. (2007). Modelling soccer matches using bivariate discrete distributions with general dependence structure. Stat. Neerlandica. 61, pp. 432-445.

11. McHale, I. and Scarf, P. (2011). Modelling the dependence of goals scored by opposing teams in international soccer matches. Stat. Modell. 11, pp. 219-236

12. Ong, S. H., Gupta, R. C., Ma, T. F. (2021). Bivariate Conway-Maxwell Poisson distribution with given Marginals and Correlation. Journal of Statistical Theory and Practice, 15(1).

13. Sellers, K, F., Morris, D. S. and Balakrishnan, N. (2016). Biriate Conway- Maxwell Poisson distribution: Formulation, Properties, X. and inference. Journal of Multivariate Analysis, 150 pp.152-168.

14. Wu and Yuen, K. C. (2003). A discrete-time risk model with interaction between classes of business. Insurance Mathematics and Economics, 33, pp. 117-133.

Thank you